
Routines sociales, contextes et corpus

Le cas de *à* et de *en*

Denis VIGIER

Université Lumière Lyon 2 / UMR ICAR

Résumé

Les constructions de type « x être à art. déf. y » déclenchant une interprétation dite de « routine sociale » ont donné lieu à de nombreuses études. Celles-ci ne se sont cependant intéressées qu'à la préposition « à », et ne se sont que rarement appuyées sur des données attestées en corpus. L'objectif de cet article est double : proposer une analyse contrastive des structures « x être en y » vs « x être à art. déf. y » à interprétation de routine sociale ; s'appuyer sur une enquête quantitative conduite sur la base de données Sketch Engine.

Abstract

Constructions “x être à art. déf. y” triggering an interpretation of “social routine” have given rise to numerous studies. However, these studies have only focused on the preposition à, and have rarely been based on corpus data. The aim of this article is twofold: proposing a contrastive analysis of the structures “x être en y” vs “x être à art. déf. y” with a social routine interpretation; drawing on a large-scale quantitative survey conducted on the Sketch Engine database.

Introduction

Les constructions de type « x être à art. déf. y » déclenchant une interprétation dite de « routine sociale » ont donné lieu à de nombreuses études (e.a. Furukawa 1997, 1998, 2010 ; Borillo 2001, Aurnague 2004, 2010, 2011, 2012, 2015 ; Corblin, 2011, 2012 ; Hernandez 2012) qui ont précisé, voire remodelé certains des concepts et des notions mis en place dans la publication séminale de Vandeloise (1988). Ces constructions ouvrent à une interprétation non exclusive de la relation de localisation dans laquelle cible et site interagissent *via* une activité bien ancrée dans les pratiques sociales. Ainsi l'énoncé « Pierre est à l'hôpital » invite-t-il par ex. à considérer, dans un contexte adéquat, que Pierre est hospitalisé, c'est-à-dire l'objet de soins délivrés dans le cadre de cet établissement ; de même, « Marie est au piano » ouvre à une interprétation selon laquelle elle joue du piano, etc.

Ces études ne couvrent néanmoins pas l'ensemble des constructions susceptibles d'évoquer une routine et laissent en suspens certaines questions qui mériteraient, selon nous, l'attention. Le présent travail souhaite donc étendre le champ de l'analyse dans deux directions en proposant : i) une analyse contrastive des structures « x être *en* y » et « x être à art. déf. y » à interprétation de routine sociale, ii) une étude quantitative conduite sur trois bases de données mises à disposition sur la plateforme SketchEngine¹.

Il est frappant d'observer en effet que depuis la publication de Vandeloise (1988), les auteurs cités *supra* se sont essentiellement intéressés à la construction situationnelle « x être à art. déf. y² ». Or un rapide (et très incomplet) survol des contextes accessibles *via* les concordanciers de Frantext ou de Sketch Engine fait rapidement apparaître d'autres constructions possibles mobilisant d'autres prépositions³. Il est tout aussi frappant de constater que ces mêmes études antérieures – l'article de Hernandez (2012)⁴ mis à part – relèvent pour l'essentiel d'une linguistique de l'introspection. Elles manipulent le plus souvent des énoncés hors contexte, laissant dans l'ombre le processus pourtant crucial de ce que Hernandez (2012 : 1764) nomme le « domaine mouvant des séquences décantées par l'emploi en discours ». C'est pourquoi nous proposons une étude quantitative visant à mettre en lumière certaines tendances observables dans l'usage. Nous souhaitons pouvoir ainsi répondre à deux questions qui ont pour partie sous-tendu la rédaction de cet article : i) que nous apprennent ces résultats quantitatifs quant à la manière dont les locuteurs exploitent l'instruction sémantique des prépositions abstraites *à* et *en* pour modeler et filtrer, en

-
1. <https://www.sketchengine.eu>
 2. Plus marginalement : « x aller à art. déf. N ». La construction « x être en N » est évoquée encore plus marginalement par Corblin (2012) et Aurnague (2012).
 3. Ce point sera repris *infra* sous § 1.4.
 4. Et plus marginalement Aurnague (2012, 2015).

discours et dans l'usage, les routines potentiellement associées aux noms sites entrant dans la construction ? ii) Dans quelle mesure les résultats de ces études quantitatives permettent-ils de progresser dans l'identification des prédicats les plus figés voire les plus lexicalisés ?

Après une première partie qui présentera un cursif état de l'art relatif aux études consacrées aux constructions à interprétation de routine sociale, nous présenterons le cadre général de notre analyse et ses objectifs. Enfin, dans une troisième partie, nous présenterons et commenterons les principaux résultats auxquels nous sommes parvenu.

1. Cursif état de l'art

Les études sur la construction « *x être à art. déf. y* » qui ont suivi les travaux de Vandeloise (1987, 1988) se sont souvent davantage intéressées à tels ou tels aspects de la problématique circonscrite par ce dernier, en vue de les préciser, de les approfondir voire de les retravailler en profondeur. Eu égard à nos objectifs, nous présenterons rapidement les principaux points abordés par les auteurs cités *supra*, en ménageant des développements plus importants pour les éléments que nous discuterons dans notre deuxième partie. Nous rappelons enfin que Hernandez (2012 : 1761-1764) propose un panorama de ces mêmes études que le lecteur pourra consulter avec profit.

1.1. La classe des constructions à interprétation de « routine sociale »

Dans son article de 1987 et plus encore de 1988, Vandeloise montre que les usages spatiaux statiques de la préposition *à*, examinés dans le cadre de la structure *x est à y*, obéissent à deux règles :

La première rend compte des emplois de « localisation » de la préposition, c'est-à-dire lorsque la séquence a pour objectif « de situer une cible dont la position est inconnue ou imprécise par rapport à un site connu » (1987 : 78).

x est à y si y localise x.

Mais il est possible que « lorsque des rituels sociaux sont évoqués par un site [...] la préposition *à* dépasse son rôle spatial pour marquer l'association de la cible avec la routine évoquée par le site. » (1987 : 77). En ce cas, c'est la règle suivante qui est pertinente :

x est à y si les positions de x et y sont associées dans une routine évoquée par y.

Vandeloise propose de nommer « site intégré » « un site capable de suggérer une routine dans une communauté linguistique suffisamment large » (Vandeloise 1988 : 135).

Voici quelques exemples :

- (1) Max est à la plage

En emploi décontextualisé⁵, l'interprétation qui s'impose est que Max (la cible) est situé sur une plage et interagit avec elle en participant aux diverses activités auxquelles elle est prototypiquement associée : lire ou somnoler sur le sable, bronzer, nager, faire des châteaux de sable ... Bien entendu, il est possible que dans le contexte d'occurrence où peut être employé cet énoncé en discours, Max fasse en réalité tout autre chose que ce que les *habitus sociaux* laissent prévoir : peindre, jouer du saxophone, donner un cours... En ce cas, la routine à laquelle s'associe préférentiellement la structure est écartée, au profit ou non d'une autre.

- (2) Luc est au piano

Hors contexte, Luc est conçu comme interagissant avec l'objet « piano » : il joue de la musique

- (3) Marie est au volant (= Marie conduit sa/une voiture)

Selon Vandeloise, trois caractéristiques distinguent les emplois de « routine sociale » des emplois strictement localisateurs de à (1988 : 136-143) :

- (i) les sites intégrés y revêtent un caractère générique ;
- (ii) ils déclenchent des contraintes spécifiques sur le choix de la cible ;
- (iii) la position de la cible y est plus précise.

Reprenons chacun de ces points.

Dans les séquences à site intégré avec à, y est un SN de type « *le N* » où l'article *le* possède le plus souvent selon l'auteur une valeur générique. (Vandeloise 1988 : 141). Pour convenablement évoquer une routine, le site intégré doit en effet être conçu dans sa plus grande généralité et non dans ses détails ou sa spécificité.

5. Corblin (2011 : 70) parle d'emploi « en l'absence d'autres indications contextuelles ». Comme le fait observer Hernandez (2012 : 1761), si les interactions cible-site surgissent des transactions avec le contexte et ne sauraient être conçues comme données d'avance, certaines des activités socialement instituées auxquelles le site se prête régulièrement s'avèrent plus saillantes (par exemple, cuisine → cuisiner). Ce sont ces scénarios socialement les plus saillants que nous cherchons à expliciter en évoquant des emplois qualifiés de « décontextualisés ».

Les emplois de routine sociale introduisent, en second lieu, des contraintes⁶ (fortes mais non infrangibles) sur la cible. Par ex., comme l'illustre Hernandez (2012 : 1760) à qui nous empruntons le premier exemple, selon la nature de la cible, le prédicat « être à la cuisine » déclenchera (*cf.* (5)) ou non (*cf.* (4)) une lecture de routine équivalant à « faire la cuisine » :

- (4) Afin de bien s'épanouir, votre eau de vie préférée doit éviter d'être remuée. C'est pourquoi un cellier muni d'un système anti-vibrations conservera de riches arômes. Pour bien l'utiliser, il est important de le placer dans un endroit plutôt calme. Par exemple, s'il **est à la cuisine**, on s'abstiendra d'installer cette dernière acquisition tout près d'un lave-vaisselle ou d'un réfrigérateur. (Celliers Ethier)
- (5) Poupette **était à la cuisine**, joviale et affairée, Lionel m'a ouvert en bras de chemise. Il a soigneusement extirpé d'un panier un château-laroze 1905, il a mis quelques disques et la conversation s'est un peu traînée. Poupette, les joues rouges et le tablier bleu sur le ventre, a apporté des croque-monsieur. « Mettez-vous à table sans moi », a-t-elle dit. (J.-P. Sartre, *Lettres au castor et à quelques autres*, vol. I (1926-1939))

Dans (5), la prédication « Poupette était à la cuisine » déclenche contextuellement une interprétation du type « cuisiner » – comme le confirme plus loin la phrase « Poupette [...] a apporté des croque-monsieur ». La cible est un agent [+hum] à même d'interagir avec l'environnement « cuisine » pour accomplir une activité. Tel n'est pas le cas dans (4) où la cible (objet inerte : « le cellier ») est simplement localisée par le site « la cuisine » mais n'interagit pas avec ce dernier.

Il reste que ces contraintes n'ont pas la sévérité des restrictions de sélection et s'avèrent flexibles⁷. Ainsi est-il possible de trouver des contextes où une interprétation de routine, à priori exclue « compte tenu des contextes directement imaginables » (Vandeloise 1988 : 140), devient possible dès lors qu'un contexte plausible lui est associé. (Voir à ce sujet Vandeloise (1988 : 138) pour « Léopold est à l'arbre » et Hernandez (2012 : 1775) pour « le chien est à l'école »).

La troisième caractéristique des interprétations de routine sociale selon Vandeloise est que la position de la cible y est plus précise⁸. C'est le cas lorsque la routine intégrée par le site exige une place précise pour la cible, comme dans les prédicats tels que : « être (à la fenêtre + au piano + au volant, + à table + à l'établi ...) ». Pour d'autres : « être à (l'université + l'hôpital + au tribunal) », Vandeloise semble envisager la notion de position de manière plus abstraite et plus vague, indiquant simplement que « la cible est également étroitement associée aux activités évoquées par le site » (1988 : 138).

6. Aurnague (2012 : 194) parle de « contraintes de nature ontologique » sur la cible.

7. La cible-agent [+hum] peut aussi être à la cuisine pour se restaurer et non cuisiner. Voir par ex. Hernandez (2012 : 1761), ex. (5).

8. Aurnague (2012 : 194) parle de « contraintes 'configurationnelles' ».

Nous ne développons pas davantage les analyses de Vandeloise et passons aux deux problématiques présentées ci-dessous, qui reprennent certains des points ayant donné lieu à discussion dans les travaux qui ont suivi et qui éclaireront, dans nos deuxième et troisième parties, notre propre démarche.

1.2. Valeurs de localisation et de routines : contraintes et recouvrements

Vandeloise (1988) assigne à la relation de localisation le rôle de « situer une cible dont la position est inconnue ou imprécise par rapport à un site connu » (p. 126). À cet égard, la spécification de la position du site, autrement dit « la précision avec laquelle elle est localisée dans le savoir partagé des locuteurs » joue un rôle de premier plan. Or, fait-il observer (e.a. p. 129), cette contrainte est levée dans le cas où la séquence « x être à art. déf. y » ouvre à une lecture de routine sociale puisque l'article y revêt une valeur générique. Autrement dit les emplois de *à* opérant une relation de localisation et ceux engendrant une lecture de routine sociale possèderaient des distributions distinctes.

Cette position a été remise en cause par les travaux notamment de Borillo (2001) et d'Aurnague (2004, 2012, 2015), suivis par Corblin (2012) qui montrent que l'on observe fréquemment en corpus des chevauchements entre les deux valeurs. Ainsi, dans un énoncé comme

(6) Les enfants **étaient à l'école** et le village, désert (Borillo, 2001)

le prédicat « être à l'école » active à la fois une interprétation locative (les enfants étaient localisés dans le site de l'école du village et ne se trouvaient donc pas dans les rues de ce dernier, désertes) et praxéologique (ils étaient en interaction avec le site institutionnel qu'est l'école). Comme le précise Hernandez (2012 : 1762), ce type de superposition s'explique ici du fait qu'il existe entre les diverses *facettes* (Cruse, 1986) associées à école (à savoir : [IMMEUBLE], [INSTITUTION], [PERSONNEL]...) un continuum qui conduit presque toujours à une « présomption d'interaction avec l'environnement » dès lors que la cible est animée.

Cette question des chevauchements fait surgir, on le verra, un obstacle non négligeable pour le codage sémantique des interprétations de routines observables en contexte discursif. Nous y reviendrons sous § 2.3.2.

1.3. Les scénarios des interactions cible-site évoquées par les sites intégrés et leur sélection en discours

Vandeloise (1988) semblait rejeter en dehors de la sphère de la linguistique proprement dite toute analyse précise des routines sociales associées aux sites intégrés : « La description des caractéristiques des routines associées relève

de l'ethnologie plus que de la linguistique » (138). De fait, il n'a pas cherché à décrire avec précision les scénarios praxéologiques associés aux routines évoquées par les sites intégrés ni à discuter de la possibilité d'un modèle susceptible de rendre compte de la manière dont se fixe linguistiquement la routine socialisée au signifié du nom site dans le stock lexical partagé d'une communauté donnée.

Les travaux qui ont suivi ont utilement cherché à mieux comprendre les processus praxéologiques, cognitifs et linguistiques (avec une attention plus marquée pour tels ou tels de ces processus selon les options théoriques de leurs auteurs) à même d'expliquer par quelles voies se fixent dans la langue les scénarios d'interaction cible-site associés aux NSite intégrés. Toutes, en particulier, s'accordent à considérer que chaque unité du lexique intègre dans son signifié un réseau complexe de situations de nature fondamentalement praxéologique, plus ou moins fortement enracinées dans l'esprit des locuteurs, reflets de scénarios plus ou moins saillants dans le vécu praxique et discursif de ces derniers. Suivant les auteurs, les modèles théoriques mobilisés, d'essence linguistique et cognitive, diffèrent.

Corblin (2011, 2012), après Borillo (2001), prend appui sur le modèle du lexique génératif de Pustejovsky (1995) qui intègre au signifié des unités du lexique un rôle téléique *via* une qualia-structure. Pour Corblin (2011), dans les constructions de type « Pierre être au NLieu-institution », c'est ce rôle téléique de l'entité-site combiné à ce qu'il nomme la F-identification du lieu-institution et au fonctionnement associatif de l'article défini placé après *à*, qui détermine, en l'absence d'autres indications contextuelles, l'interprétation de la routine. En vertu de ce dispositif, la séquence considérée « Pierre être au NLieu-institution » déclenchera l'interprétation suivante : Pierre est dans un lieu de type N et bénéficie de la fonction téléique associée à un lieu de ce type (Corblin 2011 : 67). Ainsi, la phrase « Pierre est à la clinique » déclenche l'interprétation : Pierre bénéficie du qualia téléique qui a présidé à la construction de cet établissement⁹, à savoir dispenser des soins etc. Ce qui n'exclut évidemment pas que le contexte spécifique puisse orienter l'interprétation différemment. Pierre peut être par ex. un soignant et non un soigné : un autre qualia est alors activé et un autre rôle impliqué. Mais comme y insiste Corblin (2011 : 70), le qualia téléique définitoire de la clinique demeure toujours disponible.

Hernandez (2012) mobilise quant à elle divers « concepts fondateurs de la sémantique cognitive » (1764), parmi lesquels les *frames* (Fillmore, 1982) et les *Modèles Cognitifs Idéalisés* de Lakoff (1987). Ainsi propose-t-elle de considérer

9. « Ces lieux, qu'ils soient ou non des artefacts, sont destinés à certaines activités particulières des humains. Un hôpital est un lieu qui sert à soigner les humains, une plage est un lieu au bord de l'eau utilisé (aujourd'hui au moins) par les humains pour des activités de loisir, etc. » (Corblin, 2011 : 67).

que « le degré de saillance perceptuelle et culturelle de certaines entités, leur fréquence d'apparition dans certaines situations tend à cristalliser des séquences qui, convoquées plus souvent, verront leur ancrage cognitif se traduire par une fixation linguistique » (1765). Voilà pourquoi certaines activités socialement instituées auxquelles le site se prête régulièrement sont plus saillantes que d'autres. Même si, insiste-t-elle, c'est toujours le contexte qui tranche souverainement, d'autres interactions pertinentes pouvant émerger de ce dernier selon les caractéristiques de la cible.

Aurnague enfin (2012), qui plaide pour une analyse en termes de continuum de lexicalisation¹⁰ entre des prédicats qu'il juge peu génériques (*être à la table de travail, à l'établi, au mur*) et d'autres très génériques (*être au lit, à table, au téléphone, à l'école, à l'hôpital...*), reconnaît comme Hernandez le rôle de la fréquence dans le conditionnement de cette généricité¹¹. Mais il le considère comme partiel, voyant dans le processus diachronique de lexicalisation des GP en *à art. déf. N* le ressort ultime du processus de fixation linguistique des routines auxquels ils sont préférentiellement associés.

L'enseignement que nous retiendrons de ce rapide excursus dans les analyses proposées par les auteurs pour rendre compte du processus complexe – praxéologique, cognitif et linguistique – d'intégration des routines sous forme de scénarios dans le signifié des NSite dits « intégrés » est le suivant : si l'on peut considérer que les prédicats qu'Aurnague (2004, 2012) décrit comme présentant un haut degré de généricité tendent plus fortement que les autres à déclencher une interprétation de routine préférée, c'est toujours le contexte qui, *in fine*, filtre souverainement tel ou tel scénario potentiellement disponible dans le signifié du site intégré (ce que Hernandez (2012) nomme « le potentiel évocateur [des] entités matérielles » (1765)). L'un des objectifs majeurs de notre étude est de progresser dans notre connaissance des filtrages préférentiellement opérés dans les usages en discours pour les structures ici étudiées.

10. Le concept de « lexicalisation » est loin d'être stabilisé dans les études diachroniques, et n'est pas toujours clairement distingué des processus de « dégrammaticalisation » voire de « grammaticalisation » (voir par ex. Prévost : 2006). Nous ne pouvons aborder de front cette question dans notre article. Nous y reviendrons néanmoins *infra*, en adoptant provisoirement le point de vue d'Aurnague (e.a. 2012). Une étude approfondie reste néanmoins à conduire sur les relations qui se tissent en diachronie entre l'émergence des routines sociales et les processus de figement, de lexicalisation et de grammaticalisation dont les prédicats concernés sont le siège.

11. « La 'généricité' de certains Nsite est partiellement conditionnée par le caractère plus ou moins habituel des configurations auxquelles ils réfèrent et la fréquence ». (Aurnague 2012 : 213)

2. Cadre général et objectifs de la présente étude

Comme dit dans notre introduction, nous nous proposons d'étendre le champ de l'analyse dans deux directions : l'ouverture à une nouvelle construction, le recours à une étude quantitative sur corpus électronique.

2.1. Extension aux constructions « x être en N »

Tout comme à, *en* est apte à entrer dans des constructions à interprétation de routine sociale. On peut le montrer en reprenant les trois caractéristiques qui, selon Vandeloise (1988 : 136), distinguent les emplois de « routine sociale » des emplois strictement localisateurs de à.

- a. Seules les cibles susceptibles de jouer un rôle dans la routine sont admises.

- (7) (Le voisin + ??mon chien + ??ma voiture) est en montagne [NSite = lieu]
- (8) (Le criminel + *le gardien + *l'aumônier¹²) est en prison [NSite = entité mixte]
- (9) (Pierre + toute la famille + *le chien + *les bagages) est/sont en voiture. [NSite = objet¹³]

- b. Le SN Site n'admet pas de qualification propre à un membre *particulier* de la catégorie

- (10) Le voisin est en (montagne + *montagne du Haut-Languedoc) [NSite = lieu]
- (11) Le criminel est en (prison + *prison des Baumettes) [NSite = EM]
- (12) Pierre + toute la famille est en (voiture + *voiture de ma mère). [NSite = objet]

- c. Le scénario évoqué par le site intégré précise la position de la cible.

Si les prédicats « être en prison » et « être en montagne » ne disent pas grand-chose de la position exacte de la cible (de même que « être à l'hôpital » par ex.), « être en voiture » en revanche nous en dit davantage (comme « être au volant » par ex.) : la cible est orientée face à la route, et se trouve soit face au volant, soit sur l'un des autres sièges.

12. On fait l'hypothèse qu'il s'agit du gardien et de l'aumônier affectés à la prison pour y remplir leur fonction. Ils ne peuvent être *en prison* que s'ils ont commis un crime ou un grave délit, situation non envisagée dans l'énoncé.

13. Nous adoptons dans cet article la distinction entre Lieu, Objet et Entité Mixte proposée par Aurnague, Hickmann & Vieu (2005).

Il ressort de l'examen de ces trois points a, b, c que les NSites dénotant un lieu, une entité mixte ou un objet i) exercent une contrainte sur la cible (cf. (7), (8), (9)), ii) que l'adjonction d'un modifieur du N diminue voire annule l'acceptabilité de la séquence (cf. (10), (11), (12)), iii) qu'en certains cas l'interprétation de routine permet de connaître la position précise de la cible vis-à-vis du site. En conséquence, il apparaît légitime, pour les occurrences examinées plus haut comme pour toute autre occurrence vérifiant ces trois propriétés, de considérer que l'on a affaire à une séquence à interprétation de routine sociale. Une telle conclusion n'a rien d'étonnant et croise sans surprise celles formulées par De Mulder & Amiot (2013 : 25-26) qui proposent de traiter les structures de type *x être en y* dans lesquelles *y* est un NSite relevant de la catégorie ontologique lieu, EM ou objet, comme des « localisations qualitatives¹⁴ » :

Dans l'interprétation qualitative, le premier élément de la relation exprimée par *en* se voit attribuer une propriété associée au nom introduit par cette préposition, ce qui est très souvent illustré par des exemples comme (1a) *Il est en prison*, qui implique non seulement que la personne désignée par le sujet se trouve à l'intérieur de la prison, mais aussi qu'il est prisonnier. (de Mulder & Amiot 2013 : 25)

Enfin, si la préposition *en* est parfaitement apte à entrer dans les constructions à interprétation de routine sociale, bien d'autres constructions le sont aussi. Les quelques exemples ci-dessous, extraits de Sketch Engine ou de Frantext, en témoignent :

- (13) Je **suis dans les boulons** depuis l'enfance ! Mon père est garagiste dans mon pays natal. (Sketch Engine)
- (14) Moi je vais vous dire, je **viens du vin** ! Je viens d'un monde complètement différent. (Sketch Engine)
- (15) Tu es marchand de vins en gros ? Mais, mon pauvre garçon, il faut lâcher ça tout de suite ! Tu ne **feras rien dans le vin** (J. Renard, Journal, 1894, Frantext).
- (16) Ça fait 32 ans que je **travaille dans le poulet**, je n'ai jamais vu ça. (Sketch Engine)

Ces structures attendent selon nous d'être étudiées.

14. On fera néanmoins observer que les conclusions auxquelles parviennent ces auteurs et qui valent en général pour les constructions *x être en y* ne permettent pas de répondre aux questions soulevées dans notre introduction, notamment à celle du filtrage préférentiel des routines observable dans l'usage.

2.2. Enjeux et apports d'une étude sur corpus

Notre objectif dans cet article consiste à analyser la concurrence entre *en* et *à* dans les constructions à interprétation de routine. Comme indiqué *supra*, il est frappant d'observer que la plupart des auteurs ayant travaillé sur ce type de construction, ont ménagé une place restreinte à des données textuelles issues de bases voire de corpus. Il ne s'agit pas ici de fétichiser les corpus mais il nous paraît que, dans notre optique, ce recours se justifie à au moins deux titres.

Si l'on peut en effet formuler raisonnablement l'hypothèse que la fréquence d'usage des prédicats à haute généricité, construits avec *à* ou *en*, tend massivement à confirmer la préférence d'interprétation routinière à laquelle la sémantique de ces prédicats lexicalisés s'associe en langue, pour les prédicats de moindre généricité en revanche, il paraît impossible de prévoir quelles sont les routines qui, parmi l'ensemble de celles potentiellement évoquées par le Site, vont ou non se voir préférées dans l'usage. Or cette question nous semble d'une grande importance dans notre perspective. Notre étude de la concurrence sémantico-pragmatique entre les deux prépositions *en* et *à* telle qu'elle se déploie et se négocie en discours au sein des constructions étudiées ne peut en effet se contenter d'une méthode fondée sur l'introspection linguistique. Celle-ci s'avèrerait impuissante à fournir des informations solides et sûres sur les préférences d'usages. Or ces préférences sont à même de nous livrer à leur tour des informations sur la manière dont les locuteurs exploitent en discours les instructions sémantico-pragmatiques délivrées par l'une et l'autre préposition.

Pour s'en convaincre, on prendra l'exemple du couple de prédicats « être à la gare » vs « être en gare » que nous avons soumis à notre enquête quantitative (voir *infra*). Comme montré par Cadiot (1997 : 92), si la préposition *à* « se prête à la dynamique constitutive propre à la localisation », elle le fait (à l'inverse de *de* ou de *dans* par ex.) sans « fractionner l'espace figuré ». « *À* » selon l'auteur, est « vecteur d'un point » (p. 200). Un des effets d'une telle instruction est qu'avec cette préposition, « l'assignation exacte du lieu concerné ne se fait pas dans le cadre d'une topologie fixe »¹⁵ (p. 220). Pour le prédicat qui nous concerne (« x être à la gare »), nous retiendrons quant à nous que « gare » y possède avant tout un rôle de repère et que, vérifiant en cela le caractère non fractionné de l'espace figuré par *à*, la cible peut parfaitement être située en dehors de la gare. Voilà pourquoi, au voyageur que dépose un taxi sur un dépose-minute situé devant le bâtiment d'une gare, son conducteur pourra déclarer : « Ça y est, vous êtes à la gare ». Les choses sont différentes pour *en*. Nous adopterons dans ce travail l'« instruction sémantique de saturation référentielle réciproque »

15. Ainsi l'auteur montre-t-il comment se négocie contextuellement une telle instruction sémantique de *à* dans un énoncé comme « J'irai chercher Guy au train » (1997 : 220), analyse à laquelle nous renvoyons le lecteur.

telle que définie par Cadiot (1997 : 193) pour cette préposition et que l'on peut grossièrement résumer comme suit : « *en* crée une image de coalescence « massive », sans vectorisation ni bornage » (p. 200). Rapportée à la structure qui nous occupe, on peut faire l'hypothèse que conformément à sa dynamique instructionnelle, *en* configurera le rapport cible-site comme une localisation beaucoup plus abstraite que celle opérée par *à*. De fait, l'absence de déterminant devant le N « gare » ôte toute autonomie référentielle à l'« entité mixte » (Aurnague 2004) « gare », de sorte qu'elle se voit configurée par *en* comme le « domaine de manifestation du sujet en situation » (Cadiot 1997 : 192). Cette relation de recouvrement – de coalescence massive – cible-site avec *en* qui définit des états liés « à quelque activité/contrôle » (*ibid.*) de la cible a fait l'objet de très nombreuses analyses dans la littérature consacrée à cette préposition¹⁶. Pour la construction qui nous occupe, cette relation semble exclure que la cible puisse se situer, comme c'était le cas avec *à*, à proximité de la gare et non dans la gare. On imagine ainsi mal le conducteur évoqué plus haut déclarer à son passager : « Ça y est, vous êtes en gare ». C'est seulement lorsqu'il aura pénétré dans le bâtiment que cet énoncé pourra s'appliquer à ce dernier. Si l'analyse a priori des deux prédicats considérés permet donc, à la lumière de nos connaissances linguistiques accumulées sur les prépositions mises en jeu, de dessiner un horizon des possibles interprétatifs auxquelles ouvre chacune d'elles, il demeure néanmoins hors de portée de prévoir comment l'usage va préférentiellement exploiter ces possibles. Quelles routines va-t-il privilégier ? Bien sûr, on peut forger des hypothèses, avoir des intuitions, formuler des « généralisation[s] empirique[s] assez forte[s] » (Corblin 2012 : § 80), mais seule une analyse en corpus permet de les asseoir sur une base quantitative. Cela permettant en retour au linguiste de mieux cerner comment se négocient en discours les identités sémantiques des prépositions abstraites. Car comme le rappelle Blumenthal (2008 : 37), « l'usage des mots, réalité manifeste, observable et au besoin quantifiable, est susceptible de nous renseigner sur leur signification – leur nature profonde en quelque sorte ».

2.3. Choix opérés en termes de corpus et de méthodes d'analyse

2.3.1. Bases de données utilisées

Les bases de données textuelles utilisées ont été sélectionnées dans la plateforme Sketch Engine qui propose, outre la fonctionnalité « concordance », des fonctionnalités statistiques parmi lesquelles le calcul des cooccurrences auquel nous recourrons plus loin (§ 3.4.). Notre choix de cette plateforme s'explique

16. Pour plus de détails, v. Vigier (2013).

avant tout par des nécessités de fréquences minimales d'attestation des prédicats étudiés. Les trois bases de données utilisées (French Web 2010 (frWac), frTenTen17, frTenTen20) possèdent respectivement une taille de 1.628.667.738 tokens, 6.845.630.573 tokens et 17.805.103.451 tokens. Ces longueurs rendent plus plausible la possibilité d'extraire un nombre suffisant d'occurrences attestées des prédicats étudiés pour permettre des analyses quantitatives.

2.3.2. Méthodologie suivie

Nous nous sommes concentré uniquement sur certains prédicats de type SP dans lesquels les deux prépositions sont parfaitement acceptables l'une comme l'autre. Nous nous sommes en outre restreint aux configurations mettant en jeu comme NSite des noms de lieu et d'entités mixtes (sur cette distinction, voir e.a. Aurnague, Hickmann & Vieu, 2005). Nous avons donc écarté les noms d'objet pour deux raisons majeures : ceux compatibles avec *à* pour exprimer une routine sociale ne le sont que rarement avec *en*, la réciproque étant aussi vraie. Dans les rares cas où les deux prépositions peuvent alterner (par ex. *être à la machine / en machine*), le nombre d'occurrences disponibles pour le codage après filtrage sur Sketch Engine était insuffisant pour permettre une étude quantitative.

Après avoir extrait au moyen de requêtes CQL et *via* la concordance de Sketch Engine les séquences présentées *infra*, nous les avons exportées puis codées manuellement. Lors de ce codage, nous avons systématiquement opéré un filtrage des occurrences en supprimant : i) les cas d'erreur d'étiquetage, ii) les extraits issus de traductions automatiques manifestes, iii) les productions relevant d'usages du français non-hexagonal (Suisse, Belgique, Québec ...).

Notre approche est clairement corpus-based puisque les bases de données mobilisées nous permettent d'explorer des hypothèses linguistiques forgées en amont de l'exploration outillée. Concernant le codage sémantique des séquences extraites, nous avons cherché à distinguer quelle était la relation cible-site *dominante* activée en contexte : relation à dominante de localisation ou à dominante praxéologique ? Ce choix, on s'en doute, est loin d'être trivial. En effet, comme évoqué à plusieurs reprises *supra*, les cas de chevauchements sémantiques entre localisation et praxéologie sont légion pour les interprétations de routine. En certains cas, la relation de localisation était suffisamment « pure » pour ne pas poser problème :

- (17) Sauf que mes chaussures **sont à la montagne** et que je n'y vais pas tout de suite (Sketch Engine, FrTenTen17, <https://www.skipass.com/forums/sports/ski/alpin/sujet-58001-o.html>)

Si ce type de configuration n'est pas complètement rare (cibles non-humaines), les cas les plus fréquents présentent une localisation de premier plan accompagnée d'un halo plus ou moins diffus d'activité.

- (18) À Gérardmer, la nature vous entoure à 360°. Vous **êtes à la montagne** et pourtant, l'eau est partout ! Ici, le vert des forêts et le bleu de l'eau du lac se fondent en une seule palette de couleurs merveilleusement bien assortie. Au cœur des Vosges, dans une région de lacs glaciaires, Gérardmer domine un superbe écrin naturel formé par des bois de conifères qui dévalent les pentes jusqu'au bord de l'eau.
(<https://www.lebonguide.com/destinations/lorraine/vosges/gerardmer-xonrupt-longemer>)

Le contexte présente ici de nombreuses informations ayant trait à l'espace : adverbess spatiaux (*ici, partout*), SP de localisation (*au cœur de SN, dans SN, jusqu'à SN*), lexique verbal ayant trait à l'espace (*entoure, domine, dévalent...*) qui manifestent une dominante spatiale plus que praxéologique dans l'extrait. Il n'en demeure pas moins que la séquence « vous êtes à la montagne » colore aussi la cible d'une évocation d'activité, de loisirs associés à la montagne. Et cela ne serait-ce que parce que ce texte figure sur un site d'hôtellerie qui vise une clientèle de vacanciers. Dans notre codage, nous avons affecté l'étiquette « localisation » à cet énoncé, du fait du contexte à dominante spatiale.

Voici trois autres exemples, tirés cette fois des sites de type « entités mixtes » :

- (19) Les rapports détaillés complets **sont en mairie**, à la disposition de toute personne désireuse d'en prendre connaissance dans le détail. (Sketch Engine, FrWaC, http://www.henvic.fr/pages_html/bevan2004/bevanaout04.htm)
- (20) PERMIS DE CONSTRUIRE Mon architecte ne perd pas de temps. Le permis **est en mairie**, ... 2 mois d'attente, nous sommes confiants ! (Sketch Engine, FrTenTen20, <http://taca.unblog.fr/tag/non-classe/>)
- (21) M. Vincent THIEBAUT ou l'un de ses collaborateurs **sera en mairie de Batzendorf** lundi 16 décembre 2019 de 16h00 à 17h00 où il vous accueillera et répondra à vos questions. (Sketch Engine, FrTenTen20, <http://www.batzendorf.fr/?paged=2>)

Dans (19)(20), la cible est l'objet d'une activité mise en œuvre par des agents de la mairie. Dans (19), l'interprétation contextuelle invite à considérer que les « rapports détaillés » sont conservés par les services administratifs municipaux et mis à la disposition des administrés. Ces activités de conservation et de mise à disposition sont bien réelles mais, avons-nous considéré, d'intensité moindre (car plus sporadiques) que celle dont est l'objet le permis de construire dans (20). On comprend en effet que ce document doit être l'objet d'une étude spécialisée par les services de l'urbanisme de la mairie concernée. Dans (21) enfin, la cible (humaine) est agente des procès associés au prédicat « être en mairie » et explicités contextuellement par le scripteur (« vous accueillera et répondra à vos questions »). Les contraintes opérées par le site sur la cible conduisent en outre à inférer que « M. Vincent THIEBAUT » est le maire. Dans cet énoncé, la composante praxéologique associée au prédicat « être en x » est nettement plus saillante.

Récapitulons : entre (18) et (21), il existe un gradient croissant d'activité que l'on pourrait mesurer plus précisément encore en convoquant des paramètres tels que « la cible est l'agent / le patient / le siège du procès », « l'activité est continue vs épisodique » etc. Au vu de ces éléments, nous avons considéré que l'occurrence du prédicat « être en mairie » dans (19) relevait de la catégorie dominante « localisation », les deux autres de la catégorie dominante « praxéologie ».

Faute de place, nous ne pourrons davantage commenter les options prises au cours de notre codage et qui nécessiteraient pourtant d'autres développements.

3. Résultats de l'étude quantitative conduite sur des prédicats mettant en jeu les paires « x être à art. déf. y » vs « x être en y ».

Nous commencerons par présenter les emplois où les deux séquences en concurrence possèdent une interprétation préférentielle nettement distincte l'une de l'autre au point que l'on peut se demander s'il n'y a pas une forme de partage des tâches entre elles. Nous envisagerons ensuite les emplois où cette différence est moins prononcée même si elle demeure tangible, pour terminer sur les cas où les deux constructions déclenchent des interprétations préférentielles extrêmement proches.

Les tableaux simplifiés proposés *infra* ne retiennent que certaines des valeurs codées afin de ne pas surcharger l'article. Ont ainsi été sélectionnées les têtes de colonnes suivantes :

NOM	Construction (f= xx)	% d'interprétation à dominante spatiale	Nature de la cible	Nature de l'activité cible-site	% parmi les interprétations praxéologiques identifiées
-----	-------------------------	---	-----------------------	---------------------------------------	---

Dans la colonne 2, nous avons fait figurer le nombre d'occurrences¹⁷ codées (ou fréquence absolue = f) pour chaque construction. La colonne 3 exprime en pourcentage le rapport entre le nombre d'occurrences codées « à dominante spatiale » pour la construction considérée et le nombre total d'occurrences codées pour cette même construction. La colonne 6 propose un calcul différent : le pourcentage y exprime le rapport entre le nombre d'occurrences codées « à dominante praxéologique » pour la ligne considérée et le nombre total d'occurrences codées « à dominante praxéologique » pour la construction considérée.

17. Ce nombre est le plus souvent situé entre 50 et 100. Lorsque l'on est au-dessous de 50, c'est qu'aucune base ne permettait d'obtenir plus d'occurrences après filtrage manuel.

L’objectif est de permettre au lecteur de comparer plus directement (voir en particulier sous § 3.2.¹⁸) les effectifs des sous-classes relevant de la seule catégorie sémantique « praxéologique ».

3.1. Les emplois à interprétations praxéologiques nettement polarisées

3.1.1. Les noms de lieux

Les deux Nsites retenus sont : *montagne*, *mer*. L’analyse quantitative fait apparaître que chaque préposition tend fortement à privilégier dans l’usage une interprétation praxéologique de la séquence « x être prép. y ».

NOM	Construction	% d’interprétation à dominante spatiale	Nature de la cible	Nature de l’activité cible-site	% parmi les interprétations praxéologiques identifiées
Montagne	être en (f = 95)	38%	Humain	Pratique d’un sport	70%
	être à la (f = 103)	17%	Humain	Pratique de loisirs de vacances	99%
Mer	être en (f = 95)	3%	Humain	Naviguer	88%
	être à la (f = 82)	23%	Humain	Pratique de loisirs (plage, lecture, nage...)	55%
			Navire	Naviguer	28%

La séquence « x être à la montagne » s’impose très fortement dans les contextes de loisir (99% des interprétations) alors que « x être en montagne » tend nettement à s’imposer dans les contextes où la cible pratique une activité sportive voire de compétition (70% des interprétations praxéologiques de la séquence). Voici un exemple :

- (22) [...] de la joie qui te transformait dès que **tu¹⁹ étais en montagne**, en mouvement. De ton niveau d’extra-terrestre dont trop peu de monde avait pris la mesure.
(Sketch Engine, frTenTen17, <https://www.petzl.com/FR/fr/Sport/Actu/2017-11-8/Paroles-pour-Ueli>)

La séquence « x être prép. mer » présente elle aussi une différence nette entre les interprétations praxéologiques préférées pour chacune des deux constructions :

18. Voir par ex. les interprétations praxéologiques distinguées pour « être en piscine ».
19. Il est ici question d’Ueli Steck, Alpiniste chevronné décédé.

« *x être en mer* » privilégie (88% des cas) les configurations où la cible est un humain – marin de profession ou non – qui navigue, alors que « *x être à la mer* » apparaît majoritairement (55% des interprétations praxéologiques) dans des contextes où la cible, humaine, s'adonne aux activités de loisirs associés à la mer. Plus marginalement (28% des emplois), on observe une interprétation différant de celle avec *en*, puisque si l'activité a bien trait à la navigation, la cible est un bâtiment et non un humain. Par ex. :

- (23) Une source d'énergie électrique suffisante [...] doit être disponible en permanence pendant que le navire **est à la mer**.
(Sketch Engine, FrTenTen17, https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000032591321)

On signalera enfin que si les séquences « *être (en + à la) (mer + montagne)* » peuvent localiser prioritairement la cible (jusqu'à 38% dans le cas de « *en montagne*²⁰ »), c'est le plus souvent de manière marginale, voire très marginale pour « *être en mer* » (3 % des emplois).

3.1.2. Les noms d'entités mixtes.

Ceux pour lesquels nous avons observé une différence marquée entre les interprétations praxéologiques préférentiellement activées en contexte par *à* et *en* sont : *gare, pharmacie, cuisine*.

La lecture des trois dernières colonnes du tableau ci-dessous permet de mesurer la forte polarisation des interprétations praxéologiques associées aux constructions « *x être en N* » vs « *x être à la N* ». Il est frappant aussi d'observer que *à* tend à privilégier comme cibles des actants « non professionnels » impliqués par les scénarios associables aux sites (le voyageur, le patient, le cuisinier non professionnel) alors que « *en* » tend à privilégier des actants « spécialisés » (le train, le médicament, le cuisinier professionnel). Cette observation peut être étendue aux sites « lieux » évoqués précédemment, du moins pour les cibles humaines : le marin vs le vacancier, le sportif vs le vacancier.

20. Voici un exemple : « je roule très peu, 8 000 km par an. Mes déplacements **sont en montagne** ou sur autoroute ». (Sketch Engine, frTenTen17, http://www.affection.fr/forum-suivant_77453.htm)

NOM	Construction	% d'interprétation à dominante spatiale	Nature de la cible	Nature de l'activité cible-site	% parmi les interprétations praxéologiques identifiées
Gare	être en (f= 86)	5%	Trains, rames, wagons...	Transporter des voyageurs	79%
	être à la (f= 87)	15%	Voyageurs	Prendre le / descendre du train	71%
Pharmacie	être en (f= 98)	0%	Produit pharmaceutique	Être délivré en officine	75%
	être à la (f= 39)	13%	Client	Se faire délivrer des produits pharmaceutiques	97%
Cuisine	être en (f= 92)	0%	Cuisinier	Cuisiner dans un contexte professionnel	70%
	être à la (f= 69)	19%	Membre de la famille ou ami	Cuisiner dans un contexte familial ou amical	80%

Enfin, on observe qu'avec « en », les structures « être en (gare + pharmacie + cuisine) » présentent des pourcentages d'interprétation à dominante spatiale inférieurs à 10% et à dominante praxéologique supérieurs ou égaux à 70%. Cette forte prévalence de l'interprétation praxéologique nous conduira *infra* (§ 3.4.) à élargir notre enquête sur les SP mis en jeu en dehors de leur emploi dans la structure *x être prép. y*.

3.2. Les emplois à interprétations praxéologiques modérément polarisées

Les trois noms concernés relèvent de la catégorie des entités mixtes : *mairie*, *préfecture*, *piscine*.

NOM	Construction	% d'interprétation à dominante spatiale	Nature de la cible	Nature de l'activité cible-site	% parmi les interprétations praxéologiques identifiées
Mairie	être en (f= 81)	22%	Document	Conserver et mettre à disposition	46%
	être à la (f = 105)	57%	Élu, employé, membre associatif...	Travailler, communiquer...	41%
					79%
Préfecture	être en (f = 35)	0%	Document	Conserver et mettre à disposition	49%
	être à la (f = 63)	21 %	Élu, employé, membre associatif...	Travailler, communiquer...	40%
					44%
Piscine	être en (f = 10)	0%	Humain	Nage sportive	30%
				Nage loisir	30%
				Rééducation	20%
	être à la (f = 83)	4%	Barre radioactive	Contrôle radioactivité	20%
			Humain	Nage loisir	100%

Nous ne pourrions pas commenter en détail ces données, faute de place. Nous insisterons néanmoins sur les raisons qui nous conduisent à parler, pour ces structures, de configurations « modérément polarisées » :

- au moins l’une des constructions étudiées pour un même NSite (parfois les deux, comme pour « préfecture ») ne présente pas une prévalence forte (> 70%) d’une seule interprétation praxéologique associée. Au contraire, on y observe un partage entre plusieurs interprétations (deux voire trois) dont la fréquence est relativement proche. Pour prendre un seul exemple : alors qu’« être à la mairie » tend fortement (79%) à sélectionner l’interprétation praxéologique où la cible est un élu, un employé, un membre d’une association²¹ ... et où l’activité associée

21. La séquence « x être à la mairie » illustre selon nous les limites de la « généralisation empirique forte » formulée par Corblin (2012 : § 80). En effet, si les entités mixtes comme « hôpital » et « clinique » qu’il convoque régulièrement dans Corblin (2011, 2012) illustrent efficacement la thèse qu’il défend, les NSites « mairie » et « préfecture » posent frontalement la question de leur télécité. Sont-elles destinées à bénéficier aux usagers (comme les établissements de soins) ou

consiste à travailler au sein de l'institution républicaine qu'abrite le bâtiment « mairie », à communiquer avec la population voire avec les élus, la séquence « être en mairie » voit son interprétation praxéologique se partager entre deux pôles. Un pôle où la cible appartient à la catégorie « documents » et l'activité associée, à celle de la conservation et de la mise à disposition ; un autre pôle où le scénario est pour l'essentiel homologue à celui associé préférentiellement à la séquence « être à la mairie ».

- Enfin, dans le cas du nom « mairie » du moins, l'interprétation essentiellement localisante y est plus fréquente que dans la section précédente, aussi bien pour « à » que pour « en²² ».

Autrement dit, le partage des tâches entre « à » et « en » en termes d'orientation interprétative de nature praxéologique semble beaucoup moins évident pour ce groupe de noms d'entités mixtes jouant le rôle de site.

bien leur fonction est-elle plus abstraite, liée avant tout à la présence homogène sur le territoire de la puissance publique (locale et nationale) ? L'imprévisibilité de l'interprétation routinière associée aux structures en cause plaide selon nous pour la seconde analyse et conduit à s'interroger sur les limites explicatives des qualia téléiques dans l'analyse sémantico-pragmatique de ces constructions.

22. « Être en montagne » mis à part (38%).

3.3. Les emplois à interprétations praxéologiques
faiblement polarisées voire dénuées de polarisation

Tel est le cas des noms d’entités mixtes : *usine, caserne, hôpital, clinique*

NOM	Construction	% d'interprétation à dominante spatiale	Nature de la cible	Nature de l'activité cible-site	% parmi les interprétations praxéologiques identifiées
Usine	être en (f = 20)	0 %	Objet Personnel	En fabrication	30 %
	être à l' (f = 74)	15 %		Travailler	70 %
					98%
Caserne	être en (f = 31)	0 %	Collectif (régiment, troupe, ...) Militaires, pompiers...	Être entretenu et entraîné par l'État en vue d'intervention	19 %
	être à la (f = 83)	11 %		Être entretenu et entraîné par l'État en vue d'intervention	78 %
					95%
Hôpital	être en (f = 110)	0%	Patients	Objets de soins	92%
	être à l' (f = 187)	7%			98%
Clinique	être en (f = 113)	0%			80%
	être à la (f = 77)	0%			71%

La principale caractéristique qui réunit l’ensemble de ces structures est que l’emploi de « à » ou de « en » n’oriente pas leur interprétation praxéologique vers une valeur clairement distincte. Certes, dire de quelqu’un qu’il est « à l’usine » ou « en usine » pour stipuler qu’il y exerce sa profession, ou encore qu’il est « à la clinique » ou « en clinique » pour stipuler qu’il y reçoit des soins, c’est à chaque fois sélectionner un prédicat qui diffère en ceci que la préposition ne profile pas son régime nominal de la même manière (voir *supra* § 1.5.). La préposition à maintient fréquemment en arrière-plan une localisation fonctionnelle de la cible que « en » tend à effacer complètement en activant fortement l’interprétation praxéologique. L’exemple de « être en usine » est à cet égard éloquent : le corpus montre en effet qu’il fonctionne comme un prédicat imposant à la cible d’être ouvrier. Au niveau de granularité où nous nous sommes situé pour notre codage, cette différence n’est pas perceptible. On peut néanmoins la mettre en lumière en examinant cette occurrence tirée de Sketch Engine :

(24) Lucien voulut savoir comment papa parlait aux ouvriers quand il **était à l’usine** et papa lui montra comment il fallait s’y prendre, et sa voix était toute changée. (Sartre, *L’Enfance d’un chef*)

Selon nous, la structure concurrente « quand il était en usine » ne serait guère possible dans ce contexte, eu égard à la contrainte forte que le prédicat exerce sur la cible.

Les N sites intégrés « usine, caserne » se distinguent enfin des sites « hôpital, clinique » en ceci qu'ils autorisent avec *en* une seconde lecture praxéologique préférée, inobservée avec *à* : pour le Nsite « usine », la cible, inerte, y est l'objet d'un processus de fabrication ; pour le Nsite « caserne », la cible y est un nom collectif (« bataillon, troupe, ... ») et non un seul individu²³.

Terminons par les noms « hôpital » et « clinique ». Ils retiennent l'attention à au moins deux titres. D'abord, l'interprétation des deux séquences concurrentes s'avère extrêmement proche. Ainsi :

- (25) Prier pour toutes les personnes malades, âgées, en situation de handicap, qu'elles **soient en hôpital**, clinique, maison de retraite ou à domicile ... (Sketch Engine, FrTenTen20, <https://www.jesuites.com/prier-pour-les-personnes-en-situation-de-handicap-lintention-de-priere-du-pape-pour-decembre-2023>)
- (26) Les enfants qui **sont à l'hôpital** seront heureux de voir leurs parents plus souvent. (Sketch Engine, FrTenTen10, <http://ecthycre.edres74.ac-grenoble.fr/filjours/01-02/fevo2.htm>)

Dans les deux cas, les cibles sont « hospitalisées ». Néanmoins, « être en hôpital » se distingue de « être à l'hôpital » en ceci que « en » sélectionne massivement (dans presque 90 % des cas, contre 2% avec *à*) des sous-catégories de l'entité mixte « hôpital » : psychiatrique (/psy), de jour, de nuit... La polarisation de chacune des structures se joue donc ici non pas sur le plan général des rôles actanciels affectés à la cible et au site (dans les deux cas, un humain bénéficie des soins prodigués par l'établissement), mais sur le plan plus spécifique de la spécialisation ou non du site et des soins prodigués, et de manière corrélée, sur l'affection dont souffre la cible. Il peut apparaître surprenant que le NSite « clinique » ne donne pas lieu à une polarisation aussi forte : on observe en effet que la sous-catégorisation du NSite apparaît dans des proportions beaucoup moins contrastées : 16% avec *à*, 28% avec *en*. Ce point reste à creuser.

Avant de revenir, pour conclure, sur l'ensemble des résultats présentés et commentés dans les trois sections *supra*, nous présenterons une dernière analyse quantitative mettant en jeu cette fois le calcul des cooccurrences (cf. Lebart, Pincemin & Poudat, 2019 :113-121 pour une présentation).

23. Nous avons choisi de distinguer ce cas de figure dans notre codage car il met en lumière une caractéristique de l'usage difficilement prévisible par le seul moyen de l'introspection linguistique.

3.4. Des routines sociales déclenchées par la construction situative « x être prép. y » à celles déclenchées par le SP « prép. y »

En quels cas les praxis routinisées associées préférentiellement aux constructions situationnelles « x être (en + à art. déf.) N » s'avèrent-elles associées tout aussi étroitement aux seuls SP « (en + à art. déf.) N » ? En cas de réponse affirmative, on pourra conclure que l'interprétation préférentielle de la construction avec *être* ne fait que refléter un scénario praxéologique tout aussi fortement ancré, cognitivement et linguistiquement, dans les seuls SP mis en jeu.

Dans cette optique, nous proposons d'examiner les données dans le tableau présenté ci-dessous. Elles résultent de calculs de cooccurrences conduits dans Sketch Engine (base FrTenTen17) appliqués à un sous-ensemble des SP examinés *supra*. Il s'agit de ceux dont les emplois praxéologiques l'emportent largement (> 70% des interprétations) sur leur valeur localisatrice, marginale (< 10%). Les raisons de ce choix sont les suivantes : la forte prévalence de l'interprétation praxéologique sur l'interprétation spatiale ainsi que sa spécialisation constituent des indices forts pour considérer que le SP est avancé sur le chemin de la lexicalisation.²⁴ Le sous-ensemble des SP retenus est donc : « à (la piscine + la clinique + l'hôpital), en (caserne + clinique + cuisine + gare + hôpital + mer + pharmacie + usine) ».

Pour notre calcul, la fenêtre de cooccurrence a été paramétrée sur trois mots avant le pivot. Les collocatifs (lemmes nominaux et verbaux) présentés occupent les dix premiers rangs des collocatifs (lemmes) préférés. L'indice statistique utilisé est le rapport de log-vraisemblance (*log-likelihood ratio*). Nous avons placé **en gras** les collocatifs sémantiquement convergents avec le scénario praxéologique routinisé de la construction avec « être » correspondante²⁵.

24. Comme indiqué *supra*, nous n'abordons pas de front dans cet article – faute de place – la question de la lexicalisation (e.a. Brinton & Traugott 2015 : 96) des SP étudiés. Il reste que ces deux caractéristiques (prévalence et spécialisation du sens figuré/ de l'extension de sens) sont à intégrer dans une vue plus vaste impliquant, outre le test de l'indifférence au nombre avancé par Aurnague (2015 : 8-9), un certain nombre de tests morphologiques, syntaxiques et sémantiques à même d'identifier un figement du SP (voir e.a. Borillo 1997, Fagard & De Mulder (2007), Stosic & Fagard 2019). La question de relier ce figement à un processus de lexicalisation constitue un autre volet de l'analyse. Cette voie est défendue par Aurnague (e.a. 2012 : 215, 2015 : 8-9) que nous suivons ici.

25. Dans les cellules du tableau, le verbe « être » a été mis entre crochets et non comptabilisé puisque c'est l'interprétation préférée de la séquence dans laquelle il figure qui joue le rôle d'étalon.

SP	10 premiers collocatifs lexicaux (lemmes)	nbre coll.	SP	10 premiers collocatifs lexicaux (lemmes)	nbre coll.
en caserne (f = 749)	<i>transformer, [être], couvent, convertir, logement, reconverter, vie, loger, vivre, militaire, château</i>	5/10	en hôpital (f = 4804)	<i>transformer, internier, séjour, [être], internement, convertir, placer, hospitalisation, patient, hôpital, infirmier</i>	6/10
à la clinique (f = 8210)	<i>[être], arriver, hospitaliser, aller, emmener, (se) rendre, séjour, panique, médecin, hôpital, visite, travailler</i>	6/10	en mer (f = 78141)	<i>pêche, éolien, sauvetage, sortie, jour, disparaître, promenade, balade, rejet, terre</i>	6/10
en clinique (f = 3728)	<i>hôpital, séjour, utiliser, [être], travailler, entrer, rentrer, cabinet, pratiquer, hospitalisation, transformer</i>	4/10	en pharmacie (f = 18081)	<i>préparateur, vendre, docteur, acheter, disponible, (se) trouver, étudiant, vente, médicament, viagra</i>	7/10
en cuisine (f = 21130)	<i>utiliser, cauchemar, [être], passer²⁶, utilisation, temps, chef, indispensable, tour, heure, douer</i>	1/10	à la piscine (f = 14726)	<i>accès, aller, plage, face, gratuit, accéder, direct, natation, (se) rendre, terrasse</i>	7/10
en gare (f = 14706)	<i>train, arriver, arrêt, arrivée, entrer, arrêter, tgv, [être], entrée, voyageur, accueil</i>	8/10	en usine (f = 7793)	<i>travailler, [être], fabriquer, préfabriquer, travail, transformer, monter, assembler, régler, réaliser, fabrication.</i>	2/10
à l'hôpital (f = 77620)	<i>transporter, [être], emmener, séjour, mourir, conduire, urgence, admettre, (se) rendre, décéder, aller</i>	10/10			

L'interprétation que nous proposons de ces résultats est la suivante : plus le nombre de collocatifs (« nbr coll. ») ayant un emploi convergent avec la routine préférentielle correspondante est élevé, plus le SP est avancé sur la voie de la lexicalisation. On trouve ainsi au haut du classement (> 6/10) : « à l'hôpital », « en gare », « en pharmacie », « à la piscine », tandis que « en usine » et « en cuisine » ($\leq 2/10$) ferment la marche. On ne commentera que ces deux derniers SP.

26. Il s'agit essentiellement de l'expression « passer du temps en cuisine », collocatif (avec « temps », donc) dont le rang élevé signale l'importance que revêt la durée dans l'expérience de l'activité culinaire pour les scripteurs.

- **En usine** : la plupart des collocatifs les plus significatifs mis au jour par le calcul de cooccurrence concerne des entités inertes objets d'un processus de fabrication. Les lemmes présentés dans le tableau correspondent en effet à des participes passés de sens passif (*fabriqué, préfabriqué, transformé* etc.). Tout se passe comme si la construction situative avait fixé (contrairement à *être en gare* par ex.) une routine faiblement saillante dans les autres usages du SP.
- **En cuisine** : hormis le « biais » introduit par la série télévisée « Cauchemar en cuisine » qui confère un poids statistique inattendu au collocatif « cauchemar », la plupart des autres termes associés au pivot renvoient en contexte aussi bien à un usage professionnel que non-professionnel du site. La construction « être en N » s'est donc spécialisée dans l'un des deux scénarios disponibles, laissant à « être à la cuisine » l'autre.

Conclusion

Comme Hernandez (2012), nous considérons les routines sociales conventionnalisées potentiellement évoquées par les Nsites intégrés dans la construction « x à art. déf. y » comme des scénarios praxiques et discursifs socialement partagés à fort ancrage cognitif. La contrepartie linguistique de cet ancrage peut parfois conduire des noms de sites à s'associer préférentiellement à telle ou telle routine dans le stock mental des locuteurs (par ex. la cuisine avec l'activité de cuisiner, l'hôpital avec celle d'être l'objet de soins ...). Il demeure néanmoins i) que pour bon nombre d'entre eux, il est difficile d'identifier, en l'absence d'indices contextuels, une routine saillante qui leur serait préférentiellement associée, ii) et que dans tous les cas, c'est le contexte qui seul fait émerger et filtre les scénarios routiniers potentiellement disponibles *via* les synthèses expérientielles ancrées dans la sémantique des noms de site. Pour reprendre l'observation de Hernandez (2012 : 1761), « les interactions cible-site surgissent des transactions avec le contexte et ne sauraient être conçues comme données d'avance ».

Voilà pourquoi dans cette étude qui s'est proposée d'élargir l'analyse des constructions de type « x être prép. (dét.) y » à la préposition *en*, nous avons choisi d'asseoir nos observations sur un vaste corpus de données textuelles issues de la plateforme Sketch Engine. Seule en effet une telle enquête était apte selon nous à mettre au jour i) les interactions cible-site – souvent imprédictibles – que sélectionnent préférentiellement en discours les locuteurs, ii) la manière dont ces derniers exploitent les instructions sémantico-pragmatiques délivrées par *en* et *à* pour modeler et filtrer ces routines en discours.

Nous avons ainsi pu mettre en lumière que, suivant les sites mobilisés, la différenciation (ou « polarisation ») des interprétations praxéologiques associables respectivement à la séquence « x être à art. déf. NSite » vs « x être en NSite » s'avérerait plus ou moins remarquable dans l'usage. Avec les NSites *cuisine, gare, mer, montagne, pharmacie*, cette polarisation est forte, alors qu'avec les NSites *usine, caserne, hôpital, clinique*, elle est faible voire inexistante – quoique l'on observe parfois des différences interprétatives à un niveau de granularité plus fin. Enfin, les NSites *mairie, piscine, préfecture* occupent une position médiane : la polarisation entre les deux constructions est modérée et on observe corrélativement un fractionnement fréquent des interprétations praxéologiques pour une même construction.

Nous avons enfin proposé de mettre à profit les résultats de notre codage et ceux obtenus par des calculs de cooccurrences menés sur certains SP pris comme pivots, pour tenter de donner une assise statistique à la catégorie des SP les plus lexicalisés d'une part, à la notion de continuum entre prédicats plus ou moins lexicalisés d'autre part.

Il conviendrait désormais d'étendre davantage le champ des études en s'intéressant prioritairement à des structures mobilisant d'autres prépositions et d'autres verbes.

Bibliographie

- AURNAGUE M. (2004). *Les structures de l'espace linguistique : regards croisés sur quelques constructions spatiales du basque et du français*. Louvain : Peeters.
- AURNAGUE M. (2010). Places-repère, localisation et routines : lorsque l'analyse du nom place rejoint celle de la préposition à. *Corela* [En ligne], HS-7, URL : <http://journals.openedition.org/corela/919> ; DOI : 10.4000/corela.919
- AURNAGUE M. (2011). « Quittant tout, nous partîmes » : « quitter » et « partir » à la lumière des changements de relation locative. *Journal of French Language Studies* 21(3), 285-312.
- AURNAGUE M. (2012). Quand la routine s'installe : remarques sur les emplois de à de type « routine sociale ». *Revue Romane* 47(2), 189-218.
- AURNAGUE M. (2015). *Pourquoi l'on n'est pas 'à la flûte' comme l'on est 'au piano' : (encore) une note sur 'à' de type "routine sociale"*. Document de travail. hal-02632371f
- AURNAGUE M., HICKMANN M. & VIEU L. (2005). Les entités spatiales dans la langue : étude descriptive, formelle et expérimentale de la catégorisation. In : C. Thinus-Blanc & J. Bullier (éds). *Agir dans l'espace*. Paris : Editions de la Maison des Sciences de l'Homme, *Cognitive*, 217-232.
- BLUMENTHAL P. (2008). Combinatoire des prépositions. Approche quantitative. *Langue Française* 157, 37-51.

- BORILLO A. (1997). Aide à l'identification des prépositions composées de temps et de lieu. *Faits de langue* 9, 173-184.
- BORILLO A. (2001). La détermination et la préposition de lieu à en français. In : X. Blanco, P.-A. Buvet & Z. Gavrilidou (éds), *Détermination et formalisation*. Amsterdam, Philadelphie : John Benjamins, 85-99.
- BRINTON L. & TRAUGOTT E. (2015). *Lexicalization and Language Change*. Cambridge : Cambridge University Press.
- CADIOT P. (1997). *Les prépositions abstraites en français*. Paris : Armand Colin.
- CORBLIN F. (2011). Des définis para-intensionnels : être à l'hôpital, aller à l'école. *Langue Française* 171, 55-75.
- CORBLIN F. (2012). Locus et telos : aller à l'école, être à la plage. *Corela* [en ligne] HS-12. URL : <http://journals.openedition.org/corela/2722>; DOI : <https://doi.org/10.4000/corela.2722>
- CRUSE D.A. (1986). *Lexical Semantics*. Cambridge : Cambridge University Press.
- DE MULDER W. & AMIOT D. (2013). *En* : de la préposition à la construction. *Langue française* 178, 21-39.
- FAGARD B. & DE MULDER W. (2007). La formation des prépositions complexes : grammaticalisation ou lexicalisation ? *Langue française* 156, 9-29.
- FILLMORE C. (1982). Frame semantics. In: Linguistic Society of Korea (ed.). *Linguistics in the Morning Calm*. Séoul : Hanhin, 111-137.
- FURUKAWA N. (1997). Les Glaneuses de Millet : emploi intensionnel de LE(S). *Revue de Sémantique et Pragmatique* 2, 169-181.
- FURUKAWA N. (1998). Cet objet curieux qu'on appelle l'article : emploi de l'article défini dans des environnements métalinguistiques. In : A. Englebert et al. (éds), *La ligne claire : de la linguistique à la grammaire. Mélanges offerts à Marc Wilmet*. Louvain-la-Neuve : Duculot, 47-54.
- FURUKAWA N. (2010). Article défini, son emploi « intensionnel » et énoncé tautologique. *Bulletin d'Etudes Françaises* 41, 51-71.
- HERNANDEZ P.-C. (2012). Est-on aux pizzas comme on est aux casseroles ? Sur les emplois métonymiques des syntagmes prépositionnels en à avec un nom d'objet. In : F. Neveu et al. (éds) : *Actes du Congrès Mondial de Linguistique Française, CMLF-2012*. ILF & EDP Sciences, 1759-1776.
- LAKOFF G. (1987). *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*. Chicago, Londres : University of Chicago Press.
- LEBART L., PINCEMIN B. & POUDAT C. (2019). *Analyse des données textuelles*. Québec : Presses de l'Université du Québec.
- PRÉVOST S. (2006). Grammaticalisation, lexicalisation et dégrammaticalisation : des relations complexes. *Cahiers de praxématique* 46, 121-140.
- PUSTEJOVSKY J. (1995). *The Generative Lexicon*. Cambridge (MA) : M.I.T. Press.
- STOSIC D. & FAGARD B. (2019). Les prépositions complexes en français. Pour une méthode d'identification multicritère. *Revue romane* 54 (1), 8-38.

- VANDELOISE C. (1987). La préposition *à* et le principe d'anticipation. *Langue Française* 76, 77-111.
- VANDELOISE C. (1988). Les usages statiques de la préposition *à*. *Cahiers de Lexicologie* 53, 119-148.
- VIGIER D. (2013). Sémantique de la préposition '*en*' : quelques repères. *Langue française* 178, 3-19.